Hebbian Learning of the Statistical and Geometrical Structure of Visual Input

Hebbian Learning of the Statistical and Geometrical Structure of Visual Input

James A. Bednar

1 **Abstract** Experiments on the visual system of carnivorous mammals have revealed complex relationships between the geometry and statistical properties of the visual world, and the geometry and statistical properties of the primary visual cortex. This review surveys an extensive body of modelling work that shows how a relatively simple set of general-purpose neural mechanisms can account for a large fraction of this observed relationship. The models consist of networks of simple artificial neurons with initially unspecific connections that are modified by Hebbian learning and homeostatic plasticity. Given examples of internally generated or visually evoked neural activity, this generic starting point develops into a realistic match to observations from the primary visual cortex, without requiring any vision-specific circuitry or neural properties. We show that the resulting network reflects both the geometrical and statistical structure of the input, and develops under constraints provided by the geometrical structure of the cortical and subcortical regions in the model. Specifically, the model neurons develop adult-like receptive fields and topographic maps selective for all of the major local visual features, and realistic topographically organized lateral connectivity that leads to systematic surround modulation effects depending on the geometry of both the visual input and the cortical representations. Together these results suggest that sensory cortices self-organize to capture the statistical properties of their inputs, revealing the underlying geometry using relatively simple local rules that allow them to build useful representations of the external environment.

James A. Bednar
Institute for Adaptive and Neural Computation, The University of Edinburgh, 10 Crichton St, EH8 9AB, Edinburgh, UK, e-mail: jbednar@inf.ed.ac.uk

To appear in *Neuromathematics of Vision*, Sarti and Citti, eds., Springer, 2014.

1 Introduction

Over the past half-century, experiments on the visual system of carnivorous mammals have revealed complex relationships between the geometry and statistical properties of the visual world, and the geometry and statistical properties of the visual cortex. For instance, an oriented line projected onto the back of the eye will evoke responses in a topographically mapped region of the primary visual cortex (V1), but in a discontinuous fashion grouped locally by orientation preference rather than retinotopic location. Figure 1 illustrates this mapping for V1 in a tree shrew, a primate-like species where the geometrical relationships are clearer because it lacks the fovea/periphery distinctions common to humans and other primates. Figure 1 shows that tree shrew V1 is organized much like the retina, with location on the retina mapping to corresponding locations in V1. But overlaid on this retinotopic map is an orientation map [18], with different patches of V1 neurons responding within the retinotopically mapped area, depending on the orientation of the input.

This patchy pattern of activity and orientation preference has been understood as the result of a dimension-reduction process [22, 26, 45], wherein the many dimensions in which a small patch of visual input could vary are mapped continuously onto the two-dimensional surface of the cortex. If the cortex had as many geometrical dimensions as the ways in which the input varied, this mapping could be straightforward. E.g. if the input varied only in retinal location (X,Y), a simple retinotopic map onto the cortical surface would suffice. Instead, the various combinations of retinotopic position and other features like orientation are flattened onto the cortical surface in a way that achieves good coverage of the inputs while maintaining local continuity [35, 45, 55]. Figure 2 illustrates this folding and flattening process for the case of ocular dominance (OD), with a cortical ocular dominance pattern interpreted as a two-dimensional view of preferences that cover a three-dimensional (X,Y,OD) space.

Cortical OR and OD maps illustrate geometric relationships between input and output spaces, but the relationships also take statistics into account. Specifically, the area of the cortical maps devoted to each feature value reflects the frequency of occurrence of that feature [50, 56]. Figure 3 shows examples of this phenomenon in cat visual cortex, for kittens reared with special goggles that blur non-vertical patterns. Similar effects occur for the OD map, when input from one eye is disrupted [59]. These results raise the possibility that the observed geometrical relationships could at least in part be the result of an underlying process of adaptation to the statistics of the input, where the input geometry constrains the possible input samples and the cortex organizes around the patterns seen on its inputs [35, 45, 58].

Although the basic dimension-reduction and folding idea is now widespread, the links between dimension reduction, input statistics, and the actual machinery and circuits in the visual cortex remains obscure. This review surveys results from a large family of closely related *mechanistic* models of V1 development, which show how the observed map patterns can arise from plausible approximations to the mechanisms present in the subcortical visual pathways and in V1. Unlike other models of the map patterns, the resulting systems can then process actual visual images, and

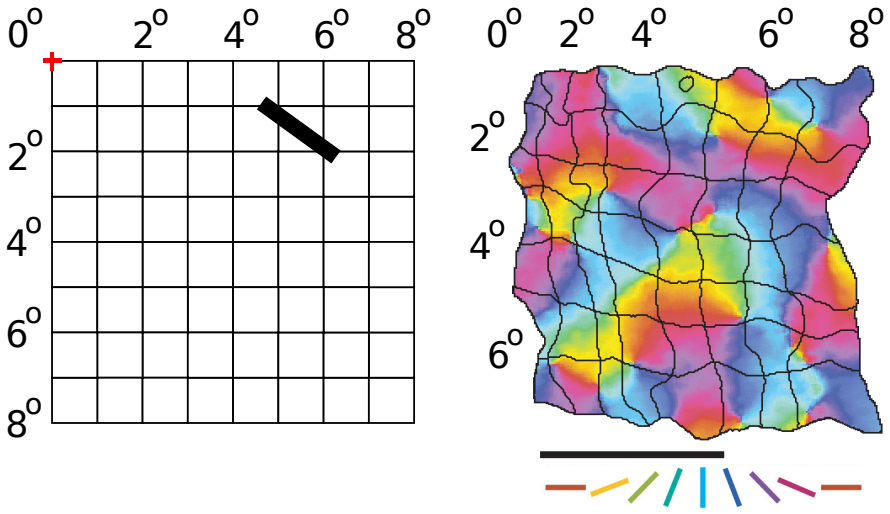
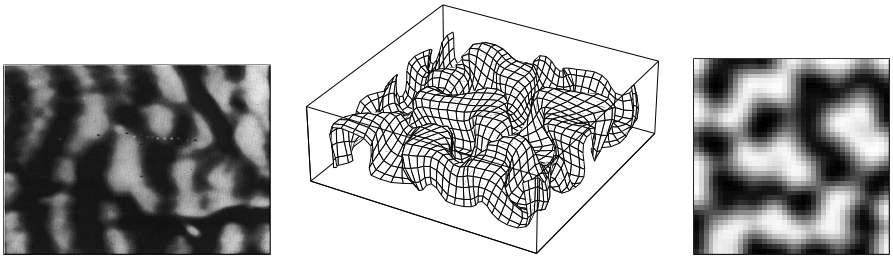


Fig. 1 Retinotopic and orientation map in V1. Given a particular fixation point (marked with a red + symbol above), the visual field seen by an animal can be divided into a regular grid, with each square representing a $1^\circ \times 1^\circ$ area of visual space. In cortical area V1 of mammals, neurons are arranged into a retinotopic map, with nearby neurons responding to nearby areas of the retina. As an example, the image on the right shows the retinotopic map on the surface of V1 of a tree shrew for an $8^\circ \times 7^\circ$ area of visual space (adapted from 21; scale bar is 1mm). A stimulus presented in a particular location in visual space (such as the thick black bar shown) evokes a response centered around the corresponding grid square in V1 (6°,2°). Which specific neurons respond within that general area, however, depends on the orientation of the stimulus. The V1 map is color coded with the preferred orientation of neurons in each location; e.g. the black bar shown at left will primarily activate neurons colored in purple in the corresponding V1 grid squares. Similar maps could be plotted for this same area showing preference for other visual features, such as motion direction, spatial frequency, color, disparity, and eye preference (depending on species). Other cortical areas are arranged into topographic maps for other sensory modalities, such as touch and audition, and for motor outputs.

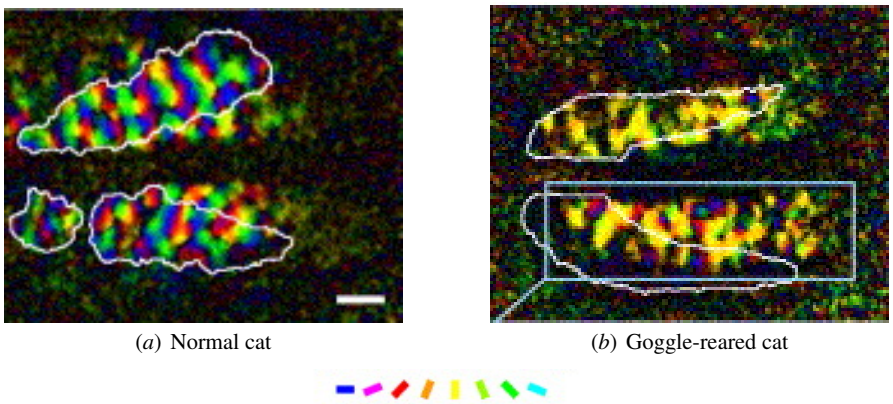
can thus be used to relate the map patterns, connectivity within the underlying networks, and observed visual and physiological phenomena. The models suggest that a wide and diverse range of observations about the visual cortex can be explained by a small set of general-purpose mechanisms. These mechanisms are not specific to vision, and should be applicable to most cortical regions.

Section 2 outlines the basic principles of these mechanistic models. Section 3) presents an implementation of a simple GCAL (Gain Control, Adaptive, Laterally connected) model [54], and describes both how it relates to earlier models on which it is based, and how it relates to a more realistic but more complex variant that covers



(*a*) Macaque OD map (*b*) Model preferences in (X,Y,OD) (*c*) Model OD

Fig. 2 Retinotopic and ocular dominance maps. (*a*) Just as for orientation, eye preference (ocular dominance; OD) is represented within the overall retinotopic map, with both of the possible eye preferences represented near any particular cortical location (here visualized across the cortical surface with black for one eye and white for the other eye); data for macaque V1 from Blasdel [17]. This pattern can be understood as a 2D projection onto the cortical surface of an underlying set of preferences in 3D: for X, Y, and ocular dominance. (*b*) shows the results of a self-organizing map (SOM) model of this organization, visualizing the 3D preference of each neuron (model from Ritter et al. 1991,1992; figure and data from Miikkulainen et al. [35]). The 2D sheet of neurons has covered the 3D input space (delineated by the box outline) by folding in the third dimension, such that every value of (X,Y,OD) is well approximated by some neuron. The resulting pattern is a type of Peano (space-filling) curve. (*c*) When the OD preference is plotted in grayscale for each neuron in their cortical locations, projecting this 3D pattern space down to the 2D cortical space, the resulting pattern is similar to animal OD maps, suggesting that animals do a similar process of representing input spaces by folding in the non-retinotopic dimensions to fill a multidimensional input space, and that cortical feature maps are the result.



(*a*) Normal cat (*b*) Goggle-reared cat

Fig. 3 Maps reflect input statistics. Comparison between orientation maps for (*a*) cats raised in a normal environment and (*b*) cats reared wearing goggles that blur non-vertical orientations shows that the distribution of orientation preferences reflects the input statistics. Thus the relationship between input and output is not merely geometrical, but is based on the statistical structure of the inputs.

all the phenomena reported here. Section 4 surveys results from GCAL and related models. Section 5 explores implications of the model, and areas for further work.

2 GCAL model overview

The models considered in this chapter are each based on the following biologically grounded principles and mechanisms (description follows [9, 10]):

1. Single-compartment (point neuron) firing-rate (i.e., non-spiking) retinal ganglion cell (RGC), lateral geniculate nucleus (LGN), and V1 model neurons (see figure 4),
2. Hardwired subcortical pathways to V1, including the main LGN or RGC cell types that have been identified,
3. Initially roughly retinotopic topographic projections from the eye to the LGN and from the LGN to V1, connecting corresponding areas of each region,
4. Initially roughly isotropic (i.e., radially uniform) local connectivity to and between neurons in layers in V1, connecting neurons non-specifically to their local and more distant neighbors,
5. Natural images and spontaneous subcortical input activity patterns that lead to V1 responses,
6. Hebbian (unsupervised activity-dependent) learning with normalization for synapses on V1 neurons,
7. Homeostatic plasticity (whole-cell adaptation of excitability to keep the average activity of each V1 neuron constant), and
8. Various modeller-determined parameters associated with each of these mechanisms, eventually intended to be set through self-regulating mechanisms.

Properties and mechanisms not necessary to explain the phenomena considered in this chapter, such as spiking, spike-timing dependent plasticity, detailed neuronal morphology, feedback from higher areas, neuromodulation, reinforcement learning, and supervised learning, have all been omitted, to clearly focus on the aspects of the system most relevant to the observed phenomena. The overall hypothesis is that much of the complex structure and properties observed in V1 emerges from interactions between relatively simple but highly interconnected computing elements, with connection strengths and patterns self-organizing in response to visual input and other sources of neural activity. Through visual experience, the geometry and statistical regularities of the visual world become encoded into the structure and connectivity of the visual cortex, leading to a complex functional cortical architecture that reflects the physical and statistical properties of the visual world.

At present, many of the results have been obtained independently in a wide variety of separate projects performed with different collaborators at different times. However, all of the models share the same underlying principles outlined above, and all are implemented using the same simulator and a small number of underlying components. See Bednar [9] for an overview of each of the different models

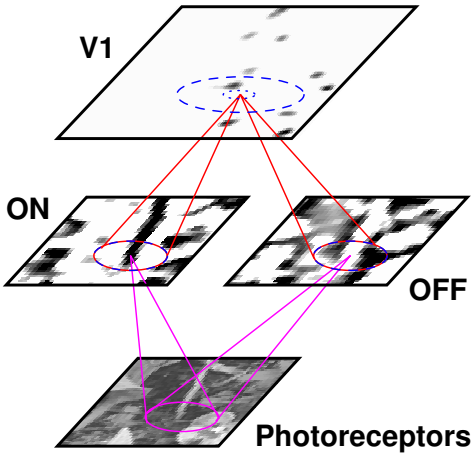


Fig. 4 Basic GCAL model architecture. In the simplest case, GCAL consists of a grayscale matrix representing the photoreceptor input, a pair of neural sheets representing the ON-center and OFF-center pathways from the photoreceptors to V1, and a single sheet representing V1. Each sheet is drawn here with a sample activity pattern resulting from one natural image patch. Each projection between sheets is illustrated with an oval showing the extent of the connection field in that projection, with lines converging on the target of the projection. Lateral projections, connecting neurons within each sheet, are marked with dashed ovals. Projections from the photoreceptors to the ON and OFF sheets, and within those sheets, are hardwired to mimic a specific class of response types found in the retina and LGN, in this case monochromatic center-surround neurons with a fixed spatial extent. Connections to and between V1 neurons adapt via Hebbian learning, allowing initially unselective V1 neurons to exhibit the range of response types seen experimentally, by differentially weighting each of the subcortical inputs (from the ON and OFF sheets) and inputs from neighboring V1 neurons. Reprinted from Bednar [10].

and how they fit together; here we present details for a simple but representative model simulating the development of orientation preferences and orientation maps for a single eye (figure 4), and describe a more complex but still incomplete "unified model" [9] covering the other phenomena, so far published only in separate models [4, 6, 9, 14, 41, 42].

The goal for each of these models is the same — to explain how a cortical network can start from an initially undifferentiated state, to wire itself into a collection of neurons that behave, at a first approximation, like those in V1. Because such a model starts with no specializations (at the cortical level) specific to vision and would organize very differently when given different inputs, it also represents a general explanation for the development and function of any sensory or motor area in the cortex.

3 GCAL Architecture

All of the models whose results are presented here are implemented in the Topographica simulator, and are freely available along with the simulator from `www.topographica.org`. Both the basic and unified models are described using the same equations shown below, previously presented in refs. [9, 54]. The model is intended to represent the visual system of the macaque monkey, but relies on data from studies of cats, ferrets, tree shrews, or other mammalian species where clear results are not yet available from monkeys.

3.1 Sheets and projections

Each Topographica model consists of a set of sheets of neurons and projections (sets of topographically mapped connections) between them. A model has sheets representing the visual input (as a set of activations in photoreceptor cells), sheets implementing the transformation from the photoreceptors to inputs driving V1 (expressed as a set of ON and OFF RGC/LGN cell activations), and sheets representing neurons in V1. The simple GCAL model (figure 4) has 4 such sheets, while the complete unified model described in Bednar [9] has 29, each representing different topographically organized populations of cells in a particular region.

Each sheet is implemented as a two-dimensional array of firing-rate neurons. The Topographica simulator allows parameters for sheets and projections to be specified in measurement units that are independent of the specific grid sizes used in a particular run of the simulation. To achieve this, Topographica sheets provide multiple spatial coordinate systems, called *sheet* and *matrix* coordinates. Where possible, parameters (e.g. sheet dimensions or connection radii) are expressed in sheet coordinates, expressed as if the sheet were a continuous neural field rather than a finite grid. In practice, of course, sheets are always simulated using some finite matrix of units. Each sheet has a parameter called its density, which specifies how many units (matrix elements) in the matrix correspond to a length of 1.0 in continuous sheet coordinates, which allows conversion between sheet and matrix coordinates. When sizes are scaled appropriately [11], results are independent of the density used, except at very low densities or for simulations with complex cells, where complexity increases with density [5]. Larger areas can be simulated easily [11], but require more memory and simulation time.

A projection to an $m \times m$ sheet of neurons consists of m^2 separate *connection fields*, one per target neuron, each of which is a spatially localized set of connections from the neurons in one input sheet that are near the location corresponding topographically to the target neuron. Figure 4 shows one sample connection field (CF) for each projection, visualized as an oval of the corresponding radius on the input sheet (drawn to scale), connected by a cone to the neuron on the target sheet (if different). The connections and their weights determine the specific properties of each neuron in the network, by differentially weighting inputs from neurons of

different types and/or spatial locations. Each of the specific types of sheets and projections is described in the following sections.

3.2 Images and photoreceptor sheets

The basic GCAL model (figure 4) has one input sheet, representing responses of photoreceptors of one cone class in one retina. The full unified GCAL model of all the input dimensions includes six input sheets (three different cone types in each eye; not shown here). For the full unified model, inputs were generated by choosing one calibrated-color image randomly from a database of single calibrated images, selecting a random patch within the image, a random direction of motion translation with a fixed speed (described in ref. [12]), and a random brightness difference between the two eyes (described in ref. [35]). These modifications are intended as a simple model of motion and eye differences, to allow development of direction preference, ocular dominance, disparity, and color maps, until suitable full-motion stereo calibrated-color video datasets of natural scenes are available. Simulated retinal waves can also be used as inputs, to provide initial receptive-field and map structure before eye opening, but are not required for receptive-field or map development in the model [13].

3.3 Subcortical sheets

The subcortical pathway from the photoreceptors to the LGN and then to V1 is represented as a set of hardwired subcortical cells with fixed connection fields (CFs) that determine the response properties of each cell. These cells represent the complete processing pathway to V1, including circuitry in the retina (including the retinal ganglion cells), the optic nerve, the lateral geniculate nucleus, and the optic radiations to V1. Because the focus of the model is to explain cortical development given its thalamic input, the properties of these ON/OFF cells are kept fixed throughout development, for simplicity and clarity of analysis.

Each distinct ON/OFF cell type is grouped into a separate sheet, each of which contains a topographically organized set of cells with identical properties but responding to a different topographically mapped region of the retinal photoreceptor input sheet. Figure 4 shows the two main different spatial response types used in the GCAL models illustrated here, ON (with an excitatory center) and OFF (with an excitatory surround). All of these cells have Difference-of-Gaussian (DoG) receptive fields, and thus perform edge enhancement at a particular size scale. Additional cell classes can easily be added as needed for spatial frequency (with multiple DoG sizes) or color (with separate cone types for the center and surround Gaussians) simulations.

For the ON and OFF cells in the unified model, there are multiple copies with different delays from the retina. These delays represent the different latencies in the lagged vs. non-lagged cells found in cat LGN [48, 62], and allow V1 neurons to become selective for the direction of motion. Many other sources of temporal delays would also lead to direction preferences, but have not been tested specifically.

3.4 Cortical sheets

Unless otherwise stated, the simulations reported in this chapter use only a single V1 sheet for simplicity, but in the full unified model, V1 is represented by multiple cortical sheets representing different cell types and different V1 layers [4, 9]. In this simplified version, cells make both excitatory and inhibitory connections (unlike actual V1 neurons), and all cells receive direct input from LGN cells (unlike many V1 neurons). Even so, the single-sheet V1 can demonstrate most of the phenomena described above, except for complex cells (which can be obtained by adding a separate population of cells without direct thalamic input [5]) and contrast-dependent surround modulation effects (which require separate populations of inhibitory and excitatory cells [4, 34]).

The behavior of the cortical sheet is primarily determined by the projections to and within it. Each of these projections is initially non-specific other than the initial rough topography, and becomes selective only through the process of self-organization (described below), which increases some connection weights at the expense of others.

3.5 Activation

The model is simulated in a series of discrete time steps with step size $\delta t = 0.05$ (roughly corresponding to 12.5 milliseconds of real time). At time 0.0, the first image is drawn on the retina, and the activation of each unit in each sheet is updated for the remaining 19 steps before time 1.0, when a new pattern is selected and drawn on the retina (and similarly until the last input pattern is drawn at time $10,000$). Each image patch on the retina represents one visual fixation (for natural images) or a snapshot of the relatively slowly changing spatial pattern of spontaneous activity (such as the well-documented retinal waves [63]). Thus the training process consists of a constant retinal activation, followed by recurrent processing at the LGN and cortical levels. For one input pattern, assume that the input is drawn on the photoreceptors at time t and the connection delay (constant for all projections) is defined as 0.05. Then at $t + 0.05$ the ON and OFF cells compute their responses, and at $t + 0.010$ the thalamic output is delivered to V1, where it similarly propagates recurrently through the intracortical projections to the cortical sheet(s) every 0.05 time steps. A much smaller step size of $\delta t = 0.002$ allows replication of the de-

tailed time course of responses to individual patterns [53], but this relatively coarse step size of 0.05 is more practical for simulations of long-term processes like neural development.

Images are presented to the model by activating the retinal photoreceptor units. The activation value $\Psi_{i,P}$ of unit i in photoreceptor sheet P is given by the brightness of that pixel in the training image.

For each model neuron in the other sheets, the activation value is computed based on the combined *activity contributions* to that neuron from each of the sheet's incoming projections. The activity contribution from a projection is recalculated whenever its input sheet activity changes, after the corresponding connection delay. For each unit j in a target sheet and an incoming projection p from sheet s_p, the activity contribution is computed from activations in the corresponding connection field F_{jp}. F_{jp} consists of the local neighborhood around j (for lateral connections), or the local neighborhood of the topographically mapped location of j on s_p (for a projection from another sheet); see examples in figures 4. The activity contribution C_{jp} to j from projection p is then a dot product of the relevant input with the weights in each connection field:

$$C_{jp}(t+\delta t) = \sum_{i \in F_{jp}} \eta_i(t)\omega_{ij,p} \tag{1}$$

where X_{is} is the activation of unit i on this projection's input sheet s_p, unit i is taken from the connection field F_{jp} of unit j, and $\omega_{ij,p}$ is the connection weight from i to j in that projection. Across all projections, multiple direct connections between the same pair of neurons are possible, but each projection p contains at most one connection between i and j, denoted by $\omega_{ij,p}$.

For a given cortical unit j, the activity $\eta_j(t+\delta t)$ is calculated from a rectified weighted sum of the activity contributions $C_{jp}(t+\delta t)$:

$$\eta_{jV}(t+\delta t) = f\left(\sum_p \gamma_p C_{jp}(t+\delta t)\right) \tag{2}$$

where f is a half-wave rectifying function with a variable threshold point (θ) dependent on the average activity of the unit, as described in the next subsection, and V denotes one of the cortical sheets.

Each γ_p is an arbitrary multiplier for the overall strength of connections in projection p. The γ_p values are set in the approximate range 0.5 to 3.0 for excitatory projections and -0.5 to -3.0 for inhibitory projections. For afferent connections, the γ_p value is chosen to map average V1 activation levels into the range 0 to 1.0 by convention, for ease of interconnecting and analyzing multiple sheets. For lateral and feedback connections, the γ_p values are then chosen to provide a balance between feedforward, lateral, and feedback drive, and between excitation and inhibition; these parameters are crucial for making the network operate in a useful regime.

ON/OFF neuron activity is computed similarly to equation 2, except to add divisive normalization and to fix the threshold θ at zero:

$$\eta_{jL}(t+\delta t) = f\left(\frac{\sum_p \gamma_p C_{jp}(t+\delta t)}{\gamma_S C_{jS}(t+\delta t)+k}\right) \quad (3)$$

where L stands for one of the ON/OFF sheets. Projection S here consists of a set of isotropic Gaussian-shaped lateral inhibitory connections (see equation 6, evaluated with $u = 1$), and p ranges over all the other projections to that sheet. k is a small constant to make the output well-defined for weak inputs. The divisive inhibition implements the contrast gain control mechanisms found in RGC and LGN neurons [1, 4, 19, 28].

Each time the activity is computed using equation 2 or 3, the new activity values are sent to each of the outgoing projections, where they arrive after the projection delay. The process of activity computation then begins again, with a new contribution C computed as in equation 1, leading to new activation values by equation 2 or 3. Activity thus spreads recurrently throughout the network, and can change, die out, or be strengthened, depending on the parameters.

With typical parameters that lead to realistic topographic map patterns, initially blurry patterns of afferent-driven cortical activity are sharpened into well-defined "activity bubbles" through locally cooperative and more distantly competitive lateral interactions [35]. Nearby neurons are thus influenced to respond more similarly, while more distant neurons receive net inhibition and thus learn to respond to different input patterns. The competitive interactions "sparsify" the cortical response into patches, in a process that can be compared to the explicit sparseness constraints in non-mechanistic models [30, 39], while the local facilitatory interactions encourage spatial locality so that smooth topographic maps will be developed.

As described in more detail below, the initially random weights to cortical neurons are updated in response to each input pattern, via Hebbian learning. Because the settling (sparsification) process typically leaves only small patches of the cortical neurons responding strongly, those neurons will be the ones that learn the current input pattern, while other nearby neurons will learn other input patterns, eventually covering the complete range of typical input variation. Overall, through a combination of the network dynamics that achieve sparsification along with local similarity, plus homeostatic adaptation that keeps neurons operating in a useful regime, plus Hebbian learning that leads to feature preferences, the network will learn smooth, topographic maps with good coverage of the space of input patterns, thereby developing into a functioning system for processing patterns of visual input without explicit specification or top-down control of this process.

3.6 Homeostatic adaptation

For this model, the threshold for activation of each cortical neuron is a very important quantity, because it directly determines how much the neuron will fire in response to a given input. Mammalian neurons appear to regulate such thresholds automatically, a process known as homeostatic plasticity or homeostatic adaptation

[57] (where homeostatic means to keep similar and stable). To set the threshold automatically, each neural unit j in V1 calculates a smoothed exponential average of its own activity ($\overline{\eta_j}$):

$$\overline{\eta_j}(t) = (1-\beta)\eta_j(t) + \beta\overline{\eta_j}(t-1) \tag{4}$$

The smoothing parameter ($\beta = 0.999$) determines the degree of smoothing in the calculation of the average. $\overline{\eta_j}$ is initialized to the target average V1 unit activity (μ), which for all simulations is $\overline{\eta_{jA}}(0) = \mu = 0.024$. The threshold is updated as follows:

$$\theta(t) = \theta(t-1) + \kappa(\overline{\eta_j}(t) - \mu) \tag{5}$$

where $\kappa = 0.0001$ is the homeostatic learning rate. The effect of this scaling mechanism is to bring the average activity of each V1 unit closer to the specified target. If the average activity of a V1 unit moves away from the target during training, the threshold for activation is thus automatically raised or lowered in order to bring it closer to the target.

3.7 Learning

Initial connection field weights are random within a two-dimensional Gaussian envelope. E.g., for a postsynaptic (target) neuron j located at sheet coordinate (0,0), the weight $\omega_{ij,p}$ from presynaptic unit i in projection p is:

$$\omega_{ij,p} = \frac{1}{Z_{\omega p}} u \exp\left(-\frac{x^2+y^2}{2\sigma_p^2}\right) \tag{6}$$

where (x,y) is the sheet-coordinate location of the presynaptic neuron i, u is a scalar value drawn from a uniform random distribution for the afferent and lateral inhibitory projections ($p = A, I$), σ_p determines the width of the Gaussian in sheet coordinates, and $Z_\omega p$ is a constant normalizing term that ensures that the total of all weights $\omega_{ij,p}$ to neuron j in projection p is 1.0, where all afferent projections are treated together as a single projection so that their sum total is 1.0. Weights for each projection are only defined within a specific maximum circular radius r_p; they are considered zero outside that radius.

Once per input pattern (after activity has settled), each connection weight ω_{ij} from unit i to unit j is adjusted using a simple Hebbian learning rule. (Learning could instead be performed at every simulation time step, but doing so would require significantly more computation time.) This rule results in connections that reflect correlations between the presynaptic activity and the postsynaptic response. Hebbian connection weight adjustment for unit j is dependent on the presynaptic activity η_i, the post-synaptic response η_j, and the Hebbian learning rate α:

$$\omega_{ij,p}(t) = \frac{\omega_{ij,p}(t-1) + \alpha\eta_j\eta_i}{\sum_k \left(\omega_{kj,p}(t-1) + \alpha\eta_j\eta_k\right)} \quad (7)$$

Unless it is constrained, Hebbian learning will lead to ever-increasing (and thus unstable) values of the weights. The weights are constrained using divisive post-synaptic weight normalization (denominator of equation 7), which is a simple and well understood mechanism. All afferent connection weights from ON/OFF sheets are normalized together in the model, which allows V1 neurons to become selective for any subset of the ON/OFF inputs. Weights are normalized separately for each of the other projections, to ensure that Hebbian learning does not disrupt the balance between feedforward drive, lateral and feedback excitation, and lateral and feedback inhibition. Subtractive normalization with upper and lower bounds could be used instead, but it would lead to binary weights [36, 37], which is not desirable for a firing-rate model whose connections represent averages over multiple physical connections. More biologically motivated homeostatic mechanisms for normalization such as multiplicative synaptic scaling [57] or a sliding threshold for plasticity [16] could be implemented instead, but these have not been tested so far.

Note that some of the results below use the earlier LISSOM model [35], which follows the same equations but lacks gain control and homeostatic adaptation (equivalent to setting $\gamma_S = 0$ and $k = 1$ in equation 3 and $\kappa = 0$ in equation 5). Without these automatic mechanisms, LISSOM requires the modeller to set the input strength and activation thresholds separately for each dataset and to adjust them as learning progresses. As long as these values have been set appropriately, previous LISSOM results can be treated equivalently to GCAL results, but GCAL is significantly simpler to use and describe, while being more robust to changes in the input distributions [54], so only GCAL is described here.

4 Results

In the following sections, we review a series of model results that account for anatomical, electrophysiological, imaging, psychophysical, and behavioral results from studies of experimental animals. Each of the results arises from the neural architecture and self-organizing mechanisms outlined in the previous section, operating on the statistical properties of the inputs, which reflect geometrical properties both of the world and of the visual system itself.

4.1 Maps and connection patterns

Figure 5 shows the pattern of orientation selectivity that emerges in the basic GCAL model from figure 4, whose subcortical pathway consists of a single set of non-lagged monochromatic ON and OFF LGN inputs for a single eye. This model ro-

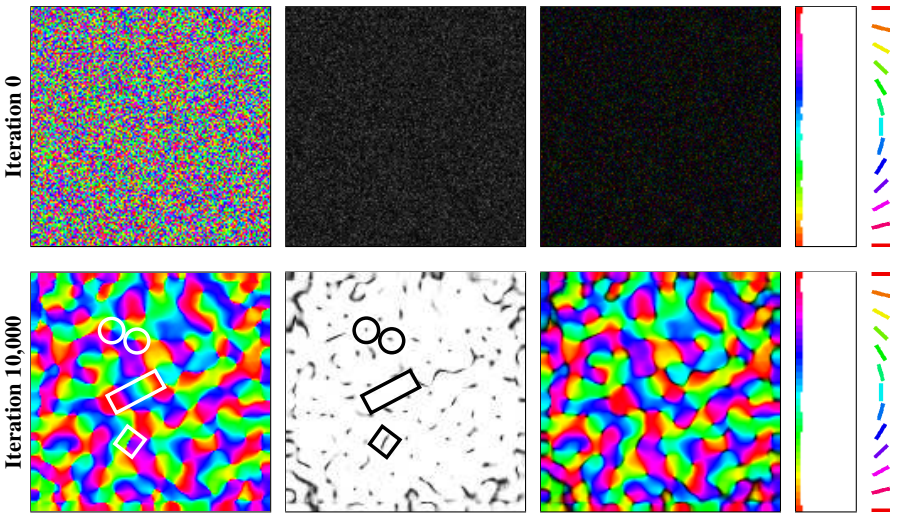


Fig. 5 Orientation maps trained on abstract stimuli (*color figure*). These plots show the orientation preference measured for each model neuron before (top row, iteration 0) and after self-organization (bottom row, iteration 10,000) based on artificially generated two-dimensional oriented Gaussian patterns. Each neuron in the map is colored according to the orientation it prefers, using color key (*e*). (*a*) The preferences are initially random (top). Through self-organization, the network developed a smoothly varying orientation map (bottom). Apart from the overall retinotopic mapping that was enforced at initialization, the map contains local geometric features found in maps from experimental animals, such as pinwheels (two are circled in white in *a* and black in *b*), linear zones (one is marked with a long white or black rectangle), and fractures (one between green and blue/purple is marked with a white or black square). (*b*) Before self-organization, the selectivity of each neuron for its (random) preferred orientation is very low (black in *b*, top). In contrast, nearly all of the self-organized neurons are highly selective for orientation (white in *b*, bottom). (*c*) Overlaying the orientation and selectivity plots shows that regions of lower selectivity in the self-organized map tend to occur near pinwheel centers and along fractures. Histograms of the number of neurons preferring each orientation are shown in (*d*), and are essentially flat because the initial weight patterns were unbiased and subsequent training inputs represented all orientations equally. These plots show that LISSOM (with GCAL getting essentially identical results as well; [54]) can develop realistic orientation maps through self-organization based on abstract input patterns. Reprinted from Bednar [8].

bustly develops orientation maps, when given training inputs that have elongated patterns. In the model, the maps emerge as the consequence of the series of activity bubbles in response to each input pattern, which cause different regions of the cortex to learn weights corresponding to different input patterns.

Over the course of development, initially unspecific connections thus become selective for specific patterns of LGN activity, including particular orientations. Hebbian learning ensures that each afferent connection field shown represents the average pattern of LGN activity that has driven that neuron to a strong response; each neuron prefers a different pattern at a specific location on the retinal surface.

Preferences from the set of all V1 neurons form a smooth topographic map covering the range of orientations present in the input patterns, yielding an orientation map similar to those from monkeys [18]. For instance, the map shows iso-feature domains, pinwheel centers, fractures, saddle points, and linear zones, with a ring-shaped Fourier transform. As in animals [49], orientation selectivity is preserved over a very wide range of contrasts, due to the effect of lateral inhibitory connections in the LGN and in V1 that normalize responses to be relative to activation of neighboring neurons rather than absolute levels of contrast [54].

Figure 6 shows that the specific map pattern observed is a consequence not of the initial random weight patterns, but of the series of randomly chosen inputs over time. The overall properties of each map will be the same for any inputs drawn from the same distribution, but the specific map pattern depends crucially on the arbitrary location and order of inputs received during self-organization. The overall type of organization primarily emerges from geometric constraints on smoothly mapping the range of values for the indicated feature, within a two-dimensional retinotopic map [32, 35].

The map patterns are also affected by the relative amount by which each feature varies in the input dataset, how often each feature appears, and other aspects of the input image statistics [13]. For instance, orientation maps trained on natural image inputs develop a preponderance of neurons with horizontal and vertical orientation preferences, as seen in ferret maps and in natural images [13, 24]. Figure 7 shows results from maps trained first on a model of spontaneous retinal activity (to account for maps present at eye opening in ferrets and cats), and then on natural images from different datasets. For natural image inputs, the map's histogram of orientation preferences will no longer be flat as it was for the artificial inputs in figure 5; instead it reflects the statistics of orientations present in the image dataset. Figure 7 shows that the model has successfully extracted the horizontal and vertical biases of natural image databases, developing many more horizontal-selective cells when trained on images with a preponderance of horizons and other horizontal patterns. This increase occurs within the context of the map already established at eye opening, with areas responding to horizontal growing larger over time, as they are activated more often than the neighboring stimuli that activate nearby regions. Figure 8 shows that the histogram of orientation preferences obtained in response to close-up natural images is a good match to that obtained for ferrets, which is intriguing because the ferrets have presumably been raised in a laboratory environment different from the forest and nature images used to train the model.

Figure 9 shows the color, motion direction, ocular dominance, spatial frequency, and disparity preferences and maps that develop when appropriate information is made available to V1 through additional ON/OFF sheets [6, 9, 41, 42]. As described in the original source for each model, the model results for each dimension have been evaluated against the available animal data, and capture the main aspects of the feature value coverage and the spatial organization of the maps [35, 41]. The maps simulated together (e.g. orientation and ocular dominance) also tend to intersect at right angles, such that high-gradient regions in one map avoid high-gradient regions in others [14]. Each neuron becomes selective for some portion of the mul-

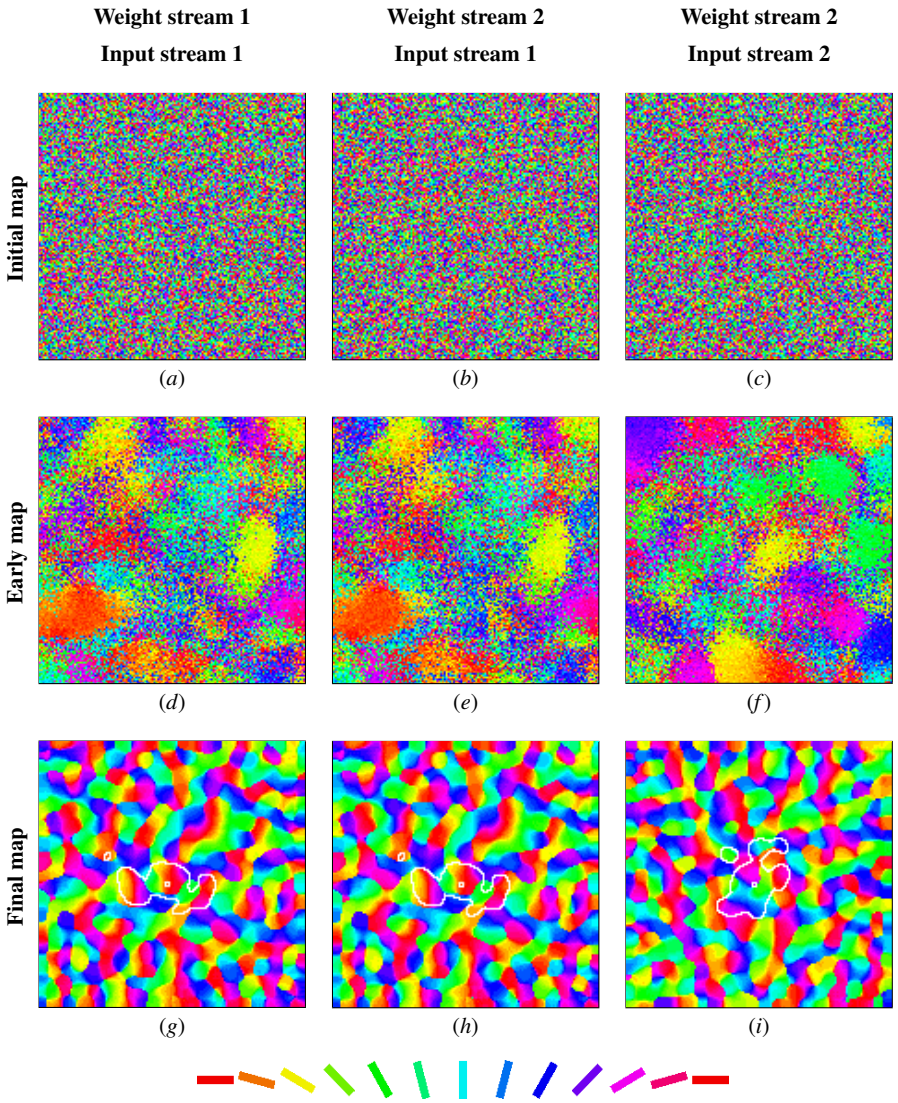


Fig. 6 Input stream determines map pattern (*color figure*). This figure shows that the self-organized orientation map patterns (e.g. in figure 5) do not depend on the random initial values of the weights. They are instead driven by the stream of input patterns presented during training. Using a different stream of random numbers for the weights results in different initial orientation maps (*a* and *b*), but has almost no effect on the final self-organized maps (compare *g* to *h*). In (*g*-*i*), the lateral inhibitory connections of one sample neuron are outlined in white, and are not affected by changing the weight stream. The final result is the same because lateral excitation smooths out differences in the initial weight values, and leads to similar large-scale patterns of activation at each iteration. (Compare maps *d* and *e* measured at iteration 100; the same large-scale features are emerging in both maps despite locally different patterns of noise caused by the different initial weights.) In contrast, changing the input stream produces very different early and final map patterns (compare *e* to *f* and *h* to *i*), even when the initial weight patterns (and therefore the initial orientation maps) are identical (*b* and *c*). Thus the input patterns are the crucial source of variation, not the initial weights. Reprinted from Bednar [8].

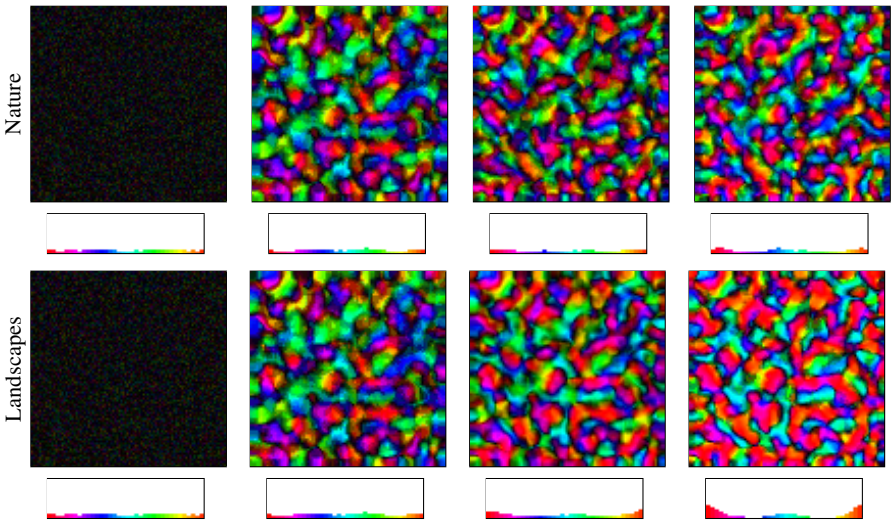
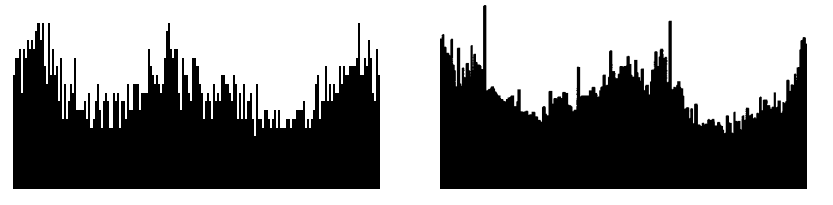


Fig. 7 Postnatal training makes orientation map match statistics of the environment (*color figure*). Each row shows results from a network trained for 1000 iterations on a model of internally generated activity [13], then trained for 9000 further iterations using natural images [51] to model postnatal visual experience. The orientation map plots (*b-d*) show selectivity as a brightness level, so that the postnatal improvement in selectivity will be visible. (*a*) and (*b*) are the same in each row. The top row shows the effect of postnatal training on natural images. With these images, more neurons become sensitive to horizontal and vertical contours, and the overall selectivity increases. However, the overall map shape remains similar, as found in laboratory animals (23; compare individual blobs between maps right to left or left to right). The postnatal changes when trained on a different database consisting primarily of landscape images are similar but much more pronounced. With these images, the network smoothly develops strong biases for vertical and horizontal contours, within the pre-determined map shape. These results show that postnatal learning can gradually adapt the prenatally developed map to match the statistics of an animal's natural environment, as shown in figure 3, while explaining how an orientation map can be present already at eye opening. Reprinted from Bednar [8].

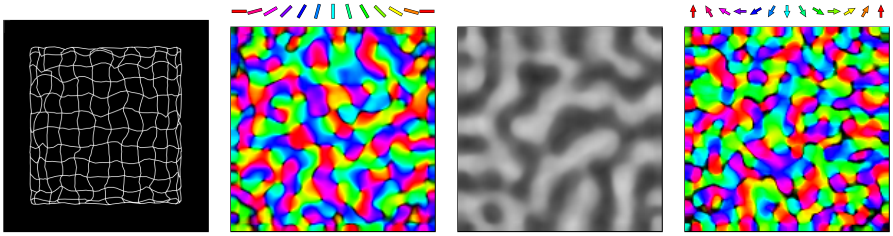
tidimensional feature space, and together they account for the variation across this space that was seen during self-organization [14].

In animals, the only large-scale information available about neural properties is from imaging techniques at the map level. In the model, it is possible to see what connectivity patterns systematically lead to the observed map preferences. Figure 10 shows these connectivity patterns for a GCAL OR map simulation with simple and complex cells, illustrating how the neurons achieve coverage of the various possible input feature values. Lateral connections, in turn, store patterns of correlation between each neuron that represent larger-scale structure and correlations. Figure 11 shows the pattern of lateral connectivity for a neuron embedded in an orientation, ocular dominance, and motion direction map. Because the lateral connections are also modified by Hebbian learning, they represent correlations between neurons, and are

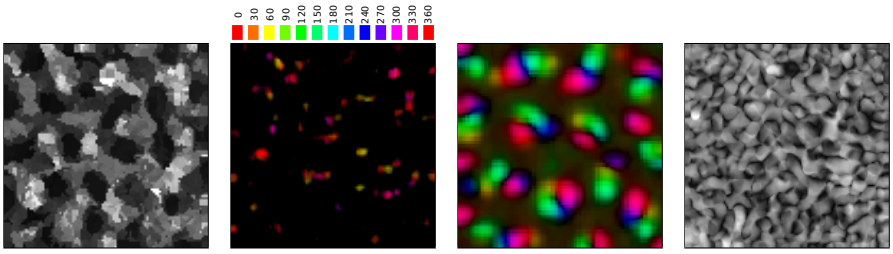


(*a*) GCAL/LISSOM model (*b*) Adult ferret

Fig. 8 Training on natural images gives matching orientation histograms. Looking more closely at the histogram for the network trained postnatally on images of natural scenes [51] shows that the resulting histograms are a close match to those found in adult ferret V1 (reprinted from 24; copyright National Academy of Sciences, U.S.A.). The model and animals both model trained on natural images have more neurons representing horizontal or vertical than oblique contours, which reflects the statistics of the natural environment. However, the natural images were chosen specifically to exclude manmade contours, while the ferrets were raised in a laboratory environment that presumably had many long edges and sharp corners, and so it may be surprising to find such a close match for these images. Work is ongoing to identify the actual pattern of first and second order statistics in natural and laboratory environments so that these results can be interpreted clearly. Reprinted from Bednar [8].

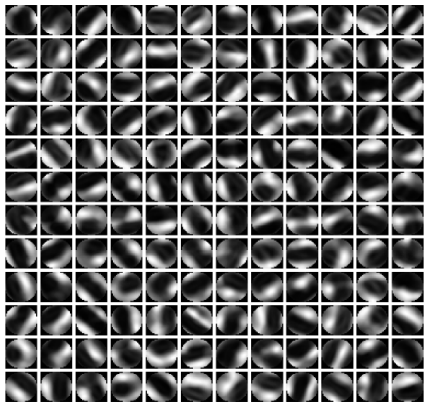


X,Y, LISSOM [35] OR, LISSOM [35] OD, LISSOM [35] DR, LISSOM [35]

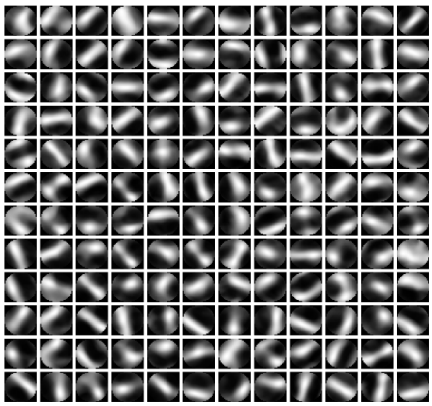


SF, LISSOM [41] DY, LISSOM [42] CR, LISSOM [6] TF, LISSOM [35]

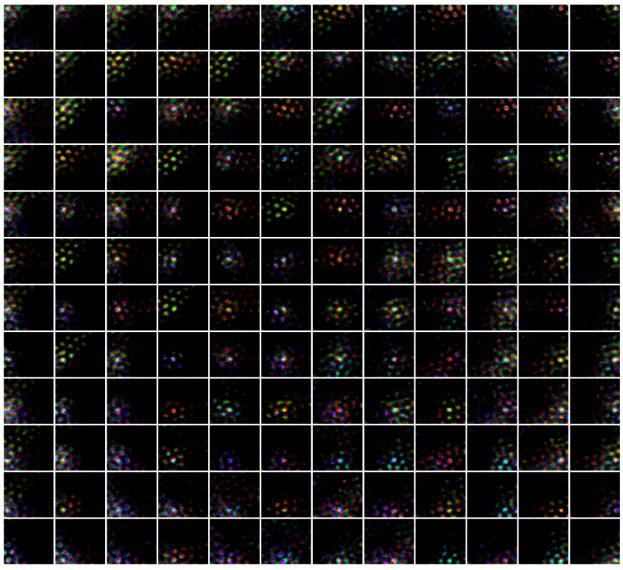
Fig. 9 Model maps for other feature dimensions. Imaging results for 4mm×4mm of model V1 from the LISSOM models of retinotopy (X,Y), orientation (OR), ocular dominance (OD), motion direction (DR), spatial frequency (SF), temporal frequency (TF), disparity (DY), and color (CR). For each spatial dimension (TF has not yet been analyzed), the model develops maps that are a close match to the experimental results. All of the maps share a property of local smoothness, which results from the short-range lateral connections in the model, but the overall patterns differ with each feature depending on how those features varied during training for that simulation. Reprinted from references indicated.



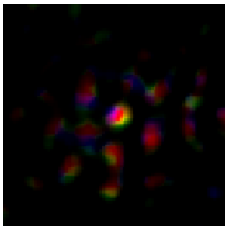
(*a*) LGN ON→V1 L4



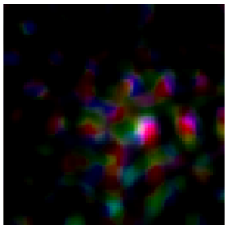
(*b*) LGN OFF→V1 L4



(*c*) V1 L2/3→V1 L2/3

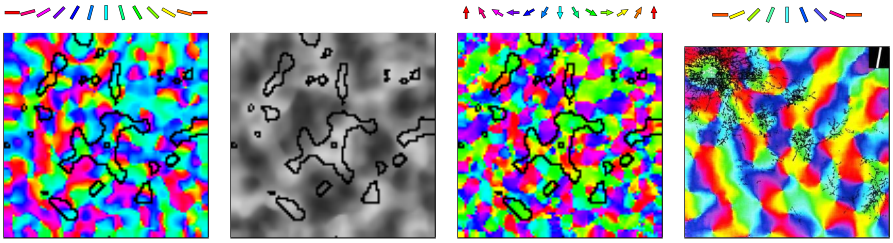


(*d*) L2/3 OR domain



(*e*) L2/3 OR pinwheel

Fig. 10 Self-organized projections to V1 L2/3. Unlike purely geometric models where the maps are represented directly in the model, the maps plotted in the previous figures are just summaries of the properties conferred on neurons by their connectivity patterns. These plots show the underlying connectivity patterns that lead to an orientation map, from a simulation with separate V1 L4 and L2/3 regions allowing the emergence of complex cells. (*a*,*b*) Connection fields from the LGN ON and OFF channels to every 20th neuron in the model L4 show that orientation preferences are reflected in the afferent connectivity to the neurons in that area. (*c*) Long-range excitatory lateral connections to those neurons preferentially come from neurons with similar OR preferences. Here strong weights are colored with the OR preference of the source neuron. Strong weights occur in clumps (appearing as small dots here) corresponding to an iso-orientation domain (each approximately 0.2–0.3mm wide); the fact that most of the dots are similar in color for any given neuron shows that the connections are orientation specific. Comparison of corresponding plots from (*c*) and (*a*) or (*b*) shows that the OR preferences of the afferent and lateral CFs are very similar. (*d*) Enlarged plot from (*c*) for a typical OR domain neuron that prefers horizontal patterns and receives connections primarily from other horizontal-preferring neurons (appearing as blobs of red or nearly red colors). (*e*) OR pinwheel neurons receive connections from neurons with many different OR preferences, because they are less selective in their responses and thus correlated with a wide range of orientation preferences. Overall, the lateral connectivity patterns reflect the patterns of co-occurrence statistics of each pair of neurons over time, due to Hebbian learning; these patterns then lead to phenomena such as orientation-specific surround modulation (figure 12). Reprinted from Antolik [4].



(*a*) OR+lateral [14] (*b*) OD+lateral [14] (*c*) DR+lateral [14] (*d*) Tree shrew; [20]

Fig. 11 Lateral connections across maps. LISSOM/GCAL neurons each participate in multiple functional maps, but have only a single set of lateral connections. Connections are strongest from other neurons with similar properties, respecting each of the maps to the degree to which that map affects correlation between neurons. Maps for a combined LISSOM OR/OD/DR simulation are shown above, with the black outlines indicating the connections to the central neuron (marked with a small black square outline) that remain after weak connections have been pruned. Model neurons receive connections from other model neurons with similar orientation preference (*a*) (as in tree shrew, (*d*)) but connections even more strongly respect the direction map (*c*). This highly monocular unit also connects strongly to the same eye (*b*), but the more typical binocular cells have wider connection distributions. Reprinted from refs. [14, 20] as indicated.

thus strong for short-range connections (due to the shared retinotopic preference of those neurons) and between other neurons often coactivated during self-organization (e.g. those sharing orientation, direction, and eye preferences). The lateral connections are thus patchy and orientation and direction specific, as found in animals [20, 47, 52]. Neurons with low levels of selectivity for any of those dimensions (e.g. binocular neurons) receive connections from a wide range of feature preferences, while highly selective neurons receive more specific connections, reflecting the different patterns of correlation in those cases. These connection patterns represent predictions, as only a few of these relationships have been tested so far in animals. The model strongly predicts that lateral connection patterns will respect all maps that account for a significant fraction of the response variance of the neurons, because each of those features will affect the correlation between neurons.

Overall, where it has been possible to make comparisons, these models have been shown to reproduce the main features of the experimental data, using a small set of assumptions. In each case, the model demonstrates how the experimentally measured map can emerge from Hebbian learning of corresponding patterns of subcortical and cortical activity. The models thus illustrate how the same basic, general-purpose adaptive mechanism will lead to very different organizations, depending on the geometrical and statistical properties of that feature. Future work will focus on showing how all the results so far could emerge simultaneously in a single model (as outlined in ref. [9]).

4.2 Surround modulation

Given a model with realistically patchy, specific lateral connectivity and realistic single-neuron properties, as described above, the patterns of interaction between neurons can be compared with neurophysiological evidence for surround modulation—influences on neural responses from distant patterns in the visual field. These studies can help validate the underlying model circuit, while helping understand how the visual cortex will respond to complicated patterns such as natural images.

For instance, figure 12 shows how the response to a sine grating patch can be modulated by a surrounding annulus. In animals, complicated patterns of interaction with the surround are seen depending on orientation and contrast [31, 50]. The model reproduces these patterns due to the orientation-specific self-organized lateral connection patterns, accounting not only for the most commonly reported and analyzed effects, but also a variety of other effects depending on the location of the neuron in the map (which affects its pattern of lateral connectivity as shown in figure 10). The model thus accounts both for the typical pattern of orientation contrast interactions, and explains why such a diversity of patterns is observed in animals. The results from these studies and related studies of size-dependent effects [4] suggest both that lateral interactions may underlie many of the observed surround modulation effects, and also that the diversity of observed effects can at least in part be traced to the diversity of lateral connection patterns, which in turn is a result of the various sequences of activations of the neurons during development.

Although the preceding results all focused on the primary visual cortex, the mechanisms involved in these models are general purpose, based only on processing statistical regularities in input patterns to reveal the underlying geometry and properties of the external world. As a demonstration, figure 13 shows that the same model can be applied to a completely non-visual input modality, rodent whiskers. The same principle of activity-bubble formation due to local cooperation and more distant competition leads to very different results for this type of input, with pinwheels that develop in a strictly aligned global organization, unlike the scattered pinwheels seen in model V1 maps. But again the results are a good match to animal data, suggesting that these general principles apply across the sensory cortex, and potentially to other cortical and subcortical regions that process patterned stimuli.

5 Discussion and future work

The results reviewed above illustrate a general approach to understanding the large-scale development, organization, and function of cortical areas, illustrating how the geometry and statistics of the external inputs interact with the geometry and architecture of the cortical architecture to determine the observed organization and operation of the visual cortex. The models show that a relatively small number of basic and largely uncontroversial assumptions and principles may be sufficient to

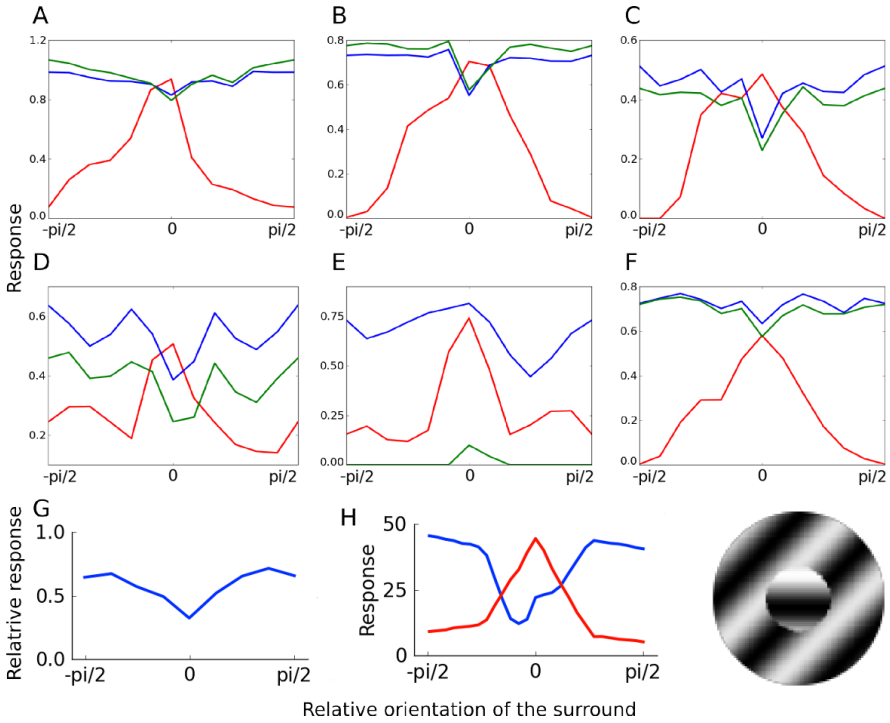
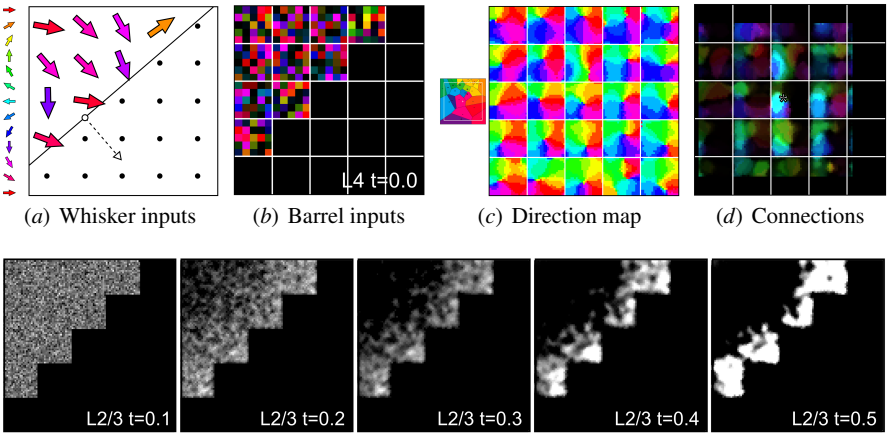


Fig. 12 Orientation-contrast tuning curves (OCTCs). For the OR model whose connection fields are shown in figure 10, the effect of the orientation-specific lateral connections can be tested using center-surround annulus stimulus like the example at the bottom right. Here the center patch is chosen to be a good match to the afferent RF of a specific V1 model neuron, and then responses are collected as the orientation of the surround is varied. In each graph A-F reprinted from ref. [4], red is the orientation tuning curve for the given neuron (with just the center grating patch), blue is for surround contrast 50%, and green is for surround contrast 100%. Top row: typically (51% of model neurons tested), a collinear surround is suppressive for these contrasts, but the surround becomes less suppressive as the surround orientation is varied (as for cat [50], G and macaque [31], H). Middle row: Other patterns seen in the model include high responses at diagonals (D, 20%, as seen in ref. [50]), strongest suppression not collinear (E, as seen in ref. [31]), and facilitation for all orientations (F, 5%). The relatively rare pattern in F has not been reported in existing studies, and thus constitutes a prediction. In each case the observed variability is a consequence of the model's Hebbian learning that leads to a diversity of patterns of lateral connectivity, rather than noise or experimental artifacts.

explain a very wide range of experimental results from the visual cortex. Even very simple neural units, i.e., firing-rate point neurons, generically connected into topographic maps with initially random or isotropic weights, can form a wide range of specific feature preferences and maps via unsupervised normalized Hebbian learning of natural images and spontaneous activity patterns. The resulting maps consist of neurons with realistic spatial response properties, with variability due to visual context and recent history that explains significant aspects of surround modulation. Combining the existing models into a single, runnable visual system is very much



(*a*) Whisker inputs (*b*) Barrel inputs (*c*) Direction map (*d*) Connections

(*e*) Time course of activation

Fig. 13 Rodent barrel cortex direction map. The V1 model in GCAL and LISSOM is completely general, and contains no vision-specific components or assumptions. As a demonstration, this figure shows results from a GCAL-based model of the rat barrel cortex, which is a primary sensory area driven by the thalamus, like V1, but with inputs ultimately from rodent whiskers instead of photoreceptors (reprinted from Wilson et al. [60]). The model has a 5×5 array of whiskers that can be deflected in any direction with different strengths; (*a*) shows a typical assumed pattern of deflection, with whiskers deflected roughly in the direction perpendicular to a moving edge (e.g. an obstacle encountered by the whiskers). The corresponding thalamic input to the barrel cortex is shown in (*b*), computed using hard-wired cosine-shaped RFs analogous to the ON and OFF channels of the LGN; bright colors indicate high activation for a unit with that direction preference. The cortical response to this pattern is initially broad (*e*), as for visual stimuli to model V1, but within a few settling iterations converges into a stable pattern of activity bubbles. Due to the geometrical arrangement of the activated whiskers, the bubbles reliably form on the leading edge of the activity pattern, which causes an immediate and strong correlation between the input patterns and the neurons that respond in barrel cortex. The result is the robust emergence of globally aligned pinwheel patterns, one per whisker barrel (*c*), which is very different from the arbitrary pinwheel patterns observed for V1 development. These patterns are a close match to experimental data from rats (2; see small map next to (*c*), showing how the map pattern emerges from the geometrical arrangement of the input stimuli and their receptors. Just as for the visual cortex models, the long-range lateral connections come from neurons with similar direction preference, due to Hebbian learning; see example for the neuron marked with a * in (*d*). For any modality, the model results reflect the geometric and statistical properties of the input, subject to constraints from the initial wiring of the cortex.

a work in progress, but the results so far suggest that doing so will be both feasible and valuable. The simulator and example simulations are freely downloadable from `www.topographica.org`, allowing any interested researcher to build on this work.

It is important to note that many of the individual results found with GCAL can also be obtained using other modelling approaches, which can be complementary to the processes modeled by GCAL. For instance, it is possible to generate orientation maps without any activity-dependent plasticity, through the initial wiring pattern

between the retina and the cortex [40, 44] or within the cortex itself [29]. Such an approach cannot explain subsequent experience-dependent development, whereas the Hebbian approach of GCAL can explain both the initial map and later plasticity, but it is of course possible that the initial map and the subsequent plasticity occur via different mechanisms. Other models are based on abstractions of some of the mechanisms in GCAL [27, 38, 61, 64], operating similarly but at a higher level. GCAL is not meant as a competitor to such models, but as a concrete, physically realizable implementation of those ideas.

As discussed throughout, the main focus of this modelling work has been on replicating experimental data using a small number of computational primitives and mechanisms, with a goal of providing a concise, concrete, and relatively simple explanation for a wide and complex range of experimental findings. A complete explanation of visual cortex development and function would go even further, demonstrating more clearly *why* the cortex should be built in this way, and precisely what information-processing purpose this circuit performs. For instance, realistic receptive fields can be obtained from "normative" models embodying the idea that the cortex is developing a set of basis functions to represent input patterns faithfully, with only a few active neurons [15, 30, 39, 43], maps can emerge by minimizing connection lengths in the cortex [33], and lateral connections can be modelled as decorrelating the input patterns [7, 25]. The GCAL model can be seen as a concrete, mechanistic implementation of these ideas, showing how a physically realizable local circuit could develop receptive fields with good coverage of the input space, via lateral interactions that also implement sparsification via decorrelation [35]. Making more explicit links between mechanistic models like GCAL and normative theories is an important goal for future work. Meanwhile, there are many aspects of cortical function not explained by current normative models. The focus of the current line of research is on first capturing those phenomena in a general-purpose mechanistic model, so that researchers can then build deeper explanations for why these computations are useful for the organism.

6 Conclusions

The GCAL model results suggest that it will soon be feasible to build a single model visual system that will account for a very large fraction of the visual response properties, at the firing rate level, of V1 neurons in a particular species. Such a model will help researchers make testable predictions to drive future experiments to understand cortical processing, as well as determine which properties require more complex approaches, such as feedback, attention, and detailed neural geometry and dynamics. The model suggests that cortical neurons develop to cover the typical range of variation in their thalamic inputs, within the context of a smooth, multidimensional topographic map, and that lateral connections store pairwise correlations and use this information to modulate responses to natural scenes, dynamically adapting to both long-term and short-term visual input statistics.

Acknowledgements Thanks to all of the collaborators whose modelling work is reviewed here, and to the members of the Institute for Adaptive and Neural Computation, and the Doctoral Training Centre in Neuroinformatics, at the University of Edinburgh, for discussions and feedback on many of the models. This work was supported in part by the UK EPSRC and BBSRC Doctoral Training Centre in Neuroinformatics, under grants EP/F500385/1 and BB/F529254/1, and by the US NIMH grant R01-MH66991. Computational resources were provided by the Edinburgh Compute and Data Facility (ECDF).

References

[1] Henry J. Alitto and W. Martin Usrey. Origin and dynamics of extraclassical suppression in the lateral geniculate nucleus of the macaque monkey. *Neuron*, 57(1):135–146, 2008.

[2] Mark L. Andermann and Christopher I. Moore. A somatotopic map of vibrissa motion direction within a barrel column. *Nature Neuroscience*, 9:543–551, 2006.

[3] James A. Anderson and E. Rosenfeld, editors. *Neurocomputing: Foundations of Research*. MIT Press, Cambridge, MA, 1988.

[4] Jan Antolik. *Unified Developmental Model of Maps, Complex Cells and Surround Modulation in the Primary Visual Cortex*. PhD thesis, School of Informatics, The University of Edinburgh, Edinburgh, UK, 2010.

[5] Jan Antolik and James A. Bednar. Development of maps of simple and complex cells in the primary visual cortex. *Frontiers in Computational Neuroscience*, 5:17, 2011.

[6] Christopher E. Ball and James A. Bednar. A self-organizing model of color, ocular dominance, and orientation selectivity in the primary visual cortex. In *Society for Neuroscience Abstracts*. Society for Neuroscience, www.sfn.org, 2009. Program No. 756.9.

[7] Horace B. Barlow and Peter Földiák. Adaptation and decorrelation in the cortex. In R. Durbin, C. Miall, and G. Mitchison, editors, *The Computing Neuron*, pages 54–72. Addison-Wesley, Reading, MA, 1989.

[8] James A. Bednar. *Learning to See: Genetic and Environmental Influences on Visual Development*. PhD thesis, Department of Computer Sciences, The University of Texas at Austin, Austin, TX, 2002. Technical Report AI-TR-02-294.

[9] James A. Bednar. Building a mechanistic model of the development and function of the primary visual cortex. *Journal of Physiology (Paris)*, 106:194–211, 2012.

[10] James A. Bednar. Constructing complex systems via activity-driven unsupervised Hebbian self-organization. In *Growing Adaptive Machines: Combining Development and Learning in Artificial Neural Networks*, Studies in Computational Intelligence. Springer, Berlin, 2013. In press.

[11] James A. Bednar, Amol Kelkar, and Risto Miikkulainen. Scaling self-organizing maps to model large cortical networks. *Neuroinformatics*, 2:275–302, 2004.

[12] James A. Bednar and Risto Miikkulainen. Self-organization of spatiotemporal receptive fields and laterally connected direction and orientation maps. *Neurocomputing*, 52–54:473–480, 2003.

[13] James A. Bednar and Risto Miikkulainen. Prenatal and postnatal development of laterally connected orientation maps. *Neurocomputing*, 58-60:985–992, 2004.

[14] James A. Bednar and Risto Miikkulainen. Joint maps for orientation, eye, and direction preference in a self-organizing model of V1. *Neurocomputing*, 69 (10–12):1272–1276, 2006.

[15] Anthony J. Bell and Terrence J. Sejnowski. The "independent components" of natural scenes are edge filters. *Vision Research*, 37:3327, 1997.

[16] E. L. Bienenstock, L. N. Cooper, and P. W. Munro. Theory for the development of neuron selectivity: Orientation specificity and binocular interaction in visual cortex. *The Journal of Neuroscience*, 2:32–48, 1982.

[17] Gary G. Blasdel. Differential imaging of ocular dominance columns and orientation selectivity in monkey striate cortex. *The Journal of Neuroscience*, 12: 3115–3138, 1992.

[18] Gary G. Blasdel. Orientation selectivity, preference, and continuity in monkey striate cortex. *The Journal of Neuroscience*, 12:3139–3161, 1992.

[19] Vincent Bonin, Valerio Mante, and Matteo Carandini. The suppressive field of neurons in lateral geniculate nucleus. *Journal of Neuroscience*, 25:10844–10856, 2005.

[20] W. H. Bosking, Y. Zhang, B. R. Schofield, and D. Fitzpatrick. Orientation selectivity and the arrangement of horizontal connections in tree shrew striate cortex. *The Journal of Neuroscience*, 17(6):2112–2127, 1997.

[21] William H. Bosking, Justin C. Crowley, and David Fitzpatrick. Spatial coding of position and orientation in primary visual cortex. *Nature Neuroscience*, 5 (9):874–882, 2002.

[22] Miguel A. Carreira-Perpiñán, Richard J. Lister, and Geoffrey J. Goodhill. A computational model for the development of multiple maps in primary visual cortex. *Cerebral Cortex*, 15(8):1222–1233, 2005.

[23] Barbara Chapman, Michael P. Stryker, and Tobias Bonhoeffer. Development of orientation preference maps in ferret primary visual cortex. *The Journal of Neuroscience*, 16(20):6443–6453, 1996.

[24] D. M. Coppola, L. E. White, D. Fitzpatrick, and D. Purves. Unequal representation of cardinal and oblique contours in ferret visual cortex. *Proceedings of the National Academy of Sciences, USA*, 95(5):2621–2623, 1998.

[25] Dawei W. Dong. Associative decorrelation dynamics: A theory of self-organization and optimization in feedback networks. In Gerald Tesauro, David S. Touretzky, and Todd K. Leen, editors, *Advances in Neural Information Processing Systems 7*, pages 925–932. Cambridge, MA: MIT Press, 1995.

[26] Richard Durbin and Graeme Mitchison. A dimension reduction framework for understanding cortical maps. *Nature*, 343:644–647, 1990.

[27] Brandon J. Farley, Hongbo Yu, Dezhe Z. Jin, and Mriganka Sur. Alteration of visual input results in a coordinated reorganization of multiple visual cortex maps. *The Journal of Neuroscience*, 27(38):10299–10310, 2007.

[28] F. Felisberti and A. M. Derrington. Long-range interactions modulate the contrast gain in the lateral geniculate nucleus of cats. *Visual Neuroscience*, 16: 943–956, 1999.

[29] Agnieszka Grabska-Barwinska and Christoph von der Malsburg. Establishment of a scaffold for orientation maps in primary visual cortex of higher mammals. *The Journal of Neuroscience*, 28(1):249–257, 2008.

[30] Aapo Hyvärinen and Patrik O. Hoyer. A two-layer sparse coding model learns simple and complex cell receptive fields and topography from natural images. *Vision Research*, 41(18):2413–2423, 2001.

[31] H. E. Jones, W. Wang, and A. M. Sillito. Spatial organization and magnitude of orientation contrast interactions in primate V1. *Journal of Neurophysiology*, 88(5):2796–2808, 2002.

[32] Matthias Kaschube, Michael Schnabel, Siegrid Löwel, David M. Coppola, Leonard E. White, and Fred Wolf. Universality in the evolution of orientation columns in the visual cortex. *Science*, 330(6007):1113–1116, 2010.

[33] Alexei A. Koulakov and Dmitri B. Chklovskii. Orientation preference patterns in mammalian visual cortex: A wire length minimization approach. *Neuron*, 29:519–527, 2001.

[34] Judith S. Law. *Modeling the Development of Organization for Orientation Preference in Primary Visual Cortex*. PhD thesis, School of Informatics, The University of Edinburgh, Edinburgh, UK, 2009.

[35] Risto Miikkulainen, James A. Bednar, Yoonsuck Choe, and Joseph Sirosh. *Computational Maps in the Visual Cortex*. Springer, Berlin, 2005.

[36] Kenneth D. Miller. A model for the development of simple cell receptive fields and the ordered arrangement of orientation columns through activity-dependent competition between ON- and OFF-center inputs. *The Journal of Neuroscience*, 14:409–441, January 1994.

[37] Kenneth D. Miller and David J. C. MacKay. The role of constraints in Hebbian learning. *Neural Computation*, 6:100–126, 1994.

[38] K. Obermayer, H. Ritter, and K. J. Schulten. A principle for the formation of the spatial structure of cortical feature maps. *Proceedings of the National Academy of Sciences, USA*, 87:8345–8349, 1990.

[39] B. A. Olshausen and D. J. Field. Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature*, 381:607–609, 1996.

[40] Se-Bum Paik and Dario L. Ringach. Retinal origin of orientation maps in visual cortex. *Nature Neuroscience*, 14(7):919–925, 2011.

[41] Christopher M. Palmer. *Topographic and Laminar Models for the Development and Organisation of Spatial Frequency and Orientation in V1*. PhD

thesis, School of Informatics, The University of Edinburgh, Edinburgh, UK, 2009.

[42] Tikesh Ramtohul. A self-organizing model of disparity maps in the primary visual cortex. Master's thesis, The University of Edinburgh, Scotland, UK, 2006.

[43] Martin Rehn and Friedrich T. Sommer. A network that uses few active neurones to code visual input predicts the diverse shapes of cortical receptive fields. *Journal of Computational Neuroscience*, 22(2):135–146, 2007.

[44] Dario L. Ringach. On the origin of the functional architecture of the cortex. *PLoS One*, 2(2):e251, 2007.

[45] Helge Ritter, Thomas Martinetz, and Klaus J. Schulten. *Neural Computation and Self-Organizing Maps: An Introduction*. Addison-Wesley, Reading, MA, 1992.

[46] Helge Ritter, K. Obermayer, Klaus J. Schulten, and Jeanne Rubner. Self-organizing maps and adaptive filters. In *Models of Neural Networks*, pages 281–306. Springer, Berlin, 1991.

[47] B. Roerig and J. P. Kao. Organization of intracortical circuits in relation to direction preference maps in ferret visual cortex. *The Journal of Neuroscience*, 19(24):RC44, 1999.

[48] A. B. Saul and A. L. Humphrey. Evidence of input from lagged cells in the lateral geniculate nucleus to simple cells in cortical area 17 of the cat. *Journal of Neurophysiology*, 68(4):1190–1208, 1992.

[49] G. Sclar and R. D. Freeman. Orientation selectivity in the cat's striate cortex is invariant with stimulus contrast. *Experimental Brain Research*, 46:457–461, 1982.

[50] F. Sengpiel, A. Sen, and C. Blakemore. Characteristics of surround inhibition in cat area 17. *Experimental Brain Research*, 116(2):216–228, 1997.

[51] H. Z. Shouval, N. Intrator, C. C. Law, and L. N. Cooper. Effect of binocular cortical misalignment on ocular dominance and orientation selectivity. *Neural Computation*, 8(5):1021–1040, 1996.

[52] Lawrence C. Sincich and Gary G. Blasdel. Oriented axon projections in primary visual cortex of the monkey. *The Journal of Neuroscience*, 21:4416–4426, 2001.

[53] Jean-Luc Stevens. A temporal model of neural activity and VSD response in the primary visual cortex. Master's thesis, The University of Edinburgh, Scotland, UK, 2011.

[54] Jean-Luc R. Stevens, Judith S. Law, Jan Antolik, and James A. Bednar. Mechanisms for stable, robust, and adaptive development of orientation maps in the primary visual cortex. *Journal of Neuroscience*, 2013. In press.

[55] Nicholas V. Swindale, D. Shoham, Amiram Grinvald, Tobias Bonhoeffer, and Mark Hubener. Visual cortex maps are optimized for uniform coverage. *Nature Neuroscience*, 3(8):822–826, 2000.

[56] Shigeru Tanaka, Jérôme Ribot, Kazuyuki Imamura, and Toshiki Tani. Orientation-restricted continuous visual exposure induces marked reorganization of orientation maps in early life. *Neuroimage*, 30(2):462–477, 2006.

[57] Gina G. Turrigiano. Homeostatic plasticity in neuronal networks: The more things change, the more they stay the same. *Trends in Neurosciences*, 22(5): 221–227, 1999.

[58] Christoph von der Malsburg. Self-organization of orientation-sensitive cells in the striate cortex. *Kybernetik*, 15:85–100, 1973. Reprinted in Anderson and Rosenfeld [3], 212–227.

[59] Torsten N. Wiesel. Postnatal development of the visual cortex and the influence of the environment. *Nature*, 299:583–591, 1982.

[60] Stuart P. Wilson, Judith S. Law, Ben Mitchinson, Tony J. Prescott, and James A. Bednar. Modeling the emergence of whisker direction maps in rat barrel cortex. *PLoS One*, 5(1):e8778, 2010.

[61] F. Wolf and T. Geisel. Universality in visual cortical pattern formation. *J Physiol Paris*, 97(2-3):253–264, 2003.

[62] Johnathan Wolfe and Larry A. Palmer. Temporal diversity in the lateral geniculate nucleus of cat. *Visual Neuroscience*, 15(4):653–675, 1998.

[63] Rachel O. L. Wong. Retinal waves and visual system development. *Annual Review of Neuroscience*, 22:29–47, 1999.

[64] Hongbo Yu, Brandon J. Farley, Dezhe Z. Jin, and Mriganka Sur. The coordinated mapping of visual space and response features in visual cortex. *Neuron*, 47(2):267–280, 2005.